Proceedings of the APAN – Research Workshop 2016
ISBN 978-4-9905448-6-7

Development of High-Precision 3D Measurement On Agriculture Using Multiple UAVs

Muhammad Haris, Seita Sukisaki, Ryo Shimomura, Zhang Heming, Li Hongyang, and Hajime Nobuhara

***Abstract*—Imaging system for high-precision 3D map on agriculture using UAVs was developed. The system was based on safe and easy UAVs with a ground station application which designed to be the interface between a human operator and the UAVs to carry out mission planning, flight command activation, and real-time flight monitoring. Based on the navigation data, and the way-points generated by the ground station, the UAVs could be automatically navigated to the desired waypoints and hover around each waypoint to collect field image data. By taking only low-resolution image, the proposed system is able to reduce the payload and increase the flight time of the UAVs. The input images then transform into higher-resolution image using reference images, taken by field server or ground-based device, via super-resolution techniques which is able to reduce blurring, blocking, and ringing artifacts especially in edge areas. Finally, we construct high-precision 3D map which proven having error of a millimeter order of magnitude. Our experiment result show that the input low-resolution can be transform into high-resolution image and effective to construct high-precision 3D map. The result indicate that the proposed system provides a reliable method of sensing agricultural field with high-precision 3D map.**

***Index Terms*—UAV, Aerial image, Sparse representation, Monitoring, Agriculture, Super-resolution, Phenotyping, 3D Images**

I. INTRODUCTION

The use of unmanned aerial vehicles (UAVs) in agriculture has increased in recent years due to problems of manual breeding methods in agriculture, which are laborious, time-consuming, unreliable, and often impossible to implement [1]–[5]. For example, high-frequency time series data are almost impossible to obtain without the use of a UAV. Moreover, large-scale, hilly landscapes make it impractical to manually analyze each tree individually using hand-held or ground-based devices. The use of UAVs can overcome such limitations, and UAV imaging offers advantages in terms of high-resolution data and precise 3D imaging.

Examples of some of the advantages offered by the use of UAVs over traditional field-based monitoring methods are listed in Table I. UAV imaging can efficiently provide high-frequency time series data, whereas aircraft and satellite systems are very complicated and their use requires arrangements be made in advance. Hand-held and ground-based devices have short preparation times but require long execution times. In terms of coverage, aircraft and satellites perform well because they can rapidly image several hectares in area, but they produce low-resolution images. In contrast, UAVs can provide better resolution as they have adjustable flight altitudes. Although hand-held and ground-based devices can provide the best resolution because they can observe parts of plants in detail, they cannot be used for large area and coverage or to produce high-frequency time series data. UAVs also require lower expenditures than aircraft or satellite as UAV sensors are much cheaper. As a UAV can be operated autonomously, control by the end user is much simpler. These advantages make UAV utilization in agricultural monitoring quite useful by offering a new perspective from which to monitor the ground with high precision [6].

TABLE I
COMPARISON OF AGRICULTURAL MONITORING SYSTEMS (◯ = SUPERIOR, △ = AVERAGE, × = POOR).

Method	Hand-held	Ground-based	UAV	Aircraft	Satellite
Frequency	×	△	◯	△	×
Coverage	×	×	△	◯	◯
Cost	◯	△	◯	×	×
User friendly	◯	△	◯	×	×
Resolution	◯	◯	△	△	×

The main problems in constructing 3D high-resolution maps using UAV images are flight-time limitations and image quality from the target object. Taking aerial images of a large field will consume a large amount of time, and to reduce time consumption, it is necessary to set an optimum height for UAV flight. However, maximizing the height, which increase the viewing perspective of the UAV and thus potentially reduces the flight time, reduces the optical detail of a target object. Therefore, it is necessary to use a super-resolution (SR) technique to obtain higher-resolution, high-precision images of target objects.

Higher-resolution image also means higher payload for the UAV, because it needs bigger bandwidth to transfer the image data to the local workstation. By using only low-resolution image from the UAV, we can reduce the payload of the UAV which also means increasing the flight time of the UAV. The input low-resolution is captured by the UAV, then downloaded

All authors are with the Computational Intelligence and Multimedia Laboratory, Department of Intelligent Interaction Technologies, University of Tsukuba, Tennoudai 1-1-1, Tsukuba 305-8573, Japan e-mail: {mharis, sukisaki, shimomura, zhang, lhy, nobuhara}@cmu.iit.tsukuba.ac.jp

Manuscript received May 1, 2016; revised June 1, 2016; accepted June 30, 2016.

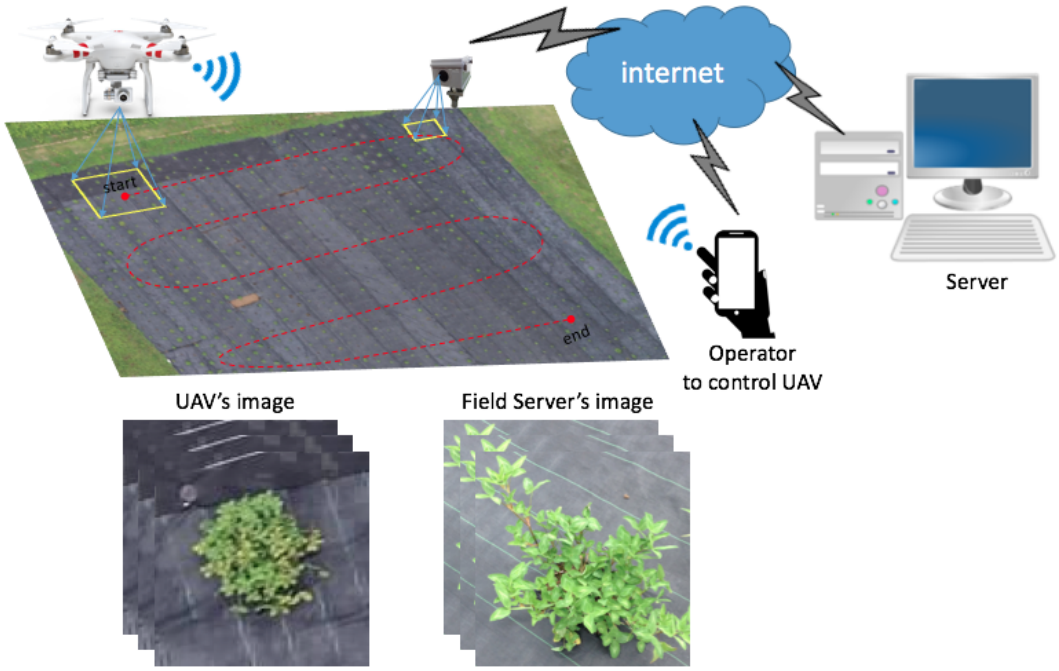

Fig. 1. Proposed system to support High-Precision Agriculture Using UAV and Field Server

to operator's smartphone and sent to server. This mechanism also enable the user to control many fields remotely. Soon after, it transform into higher-resolution using SR techniques which proposed in our paper [7]. Finally, we construct 3D image using SfM algorithm [8].

Field Server (FS) systems [9]–[12] can be used for ground-based monitoring via a series of small sensor nodes equipped with a Web server that can be accessed via the Internet and communicate, unlike traditional sensor nodes, through a wireless LAN over a high-speed transmission network. An FS system can be easily installed for remotely monitoring field information anywhere. By including the functionality of a Web server in each module, an FS system can collectively manage each module over the Internet, producing high-resolution images that can be used as training images for an SR algorithm.

In this paper, we develop a framework to construct high-precision 3D map for agriculture using UAVs and FS (or ground-based device). We propose the use of low-resolution image taken by UAVs to reduce the payload of the UAVs, and transform it into higher-resolution using proposed SR technique. Finally, we demonstrate the effectiveness of the proposed system in reconstructing 3D map. The proposed system is shown in Fig. 1.

The paper is organized as follows. Section 2 presents an explanation of the proposed system. Section 3 discusses the results of our experiments and analysis. Section 4 shows the effectiveness of the proposed system to construct 3D map. Finally, in section 5 we present our conclusions.

II. PROPOSED SYSTEM

In this section, the proposed system is explained. The proposed system have 4 main components which discussed in the following subsection. The flowchart of the proposed system is illustrated in the Fig. 2.

A. Initialize Environment

Before collecting the data, the initialization of the environment is necessary, including setting the spatial resolution, formation path, and creating waypoints that considered altitude, latitude, longitude, the distance of every turning point, flight speed, and etc. Spatial resolution was the first variable to be considered because it largely determines how much detail can be interpreted from the final image and how many images need to be obtained. Spatial resolution can be defined as how much area is represented by a pixel on the image sensor. The spatial resolution used must be large or small enough to meet the objectives of the application requirements. There are two major factors that influence spatial resolution. One is flying height and the other is the focal length of the sensor. The relationship between these three variables can be expressed as following equation [13].

$$Res = \frac{S_{pixel} \times H_f}{f} \tag{1}$$

Where Res is the spatial resolution; S_{pixel} is the pixel size of the image sensor; H_f is the flying height and f is the lens focal length. The ratio of f/H_f can be defined as image scale, which is the distance between two points on an image to the actual distance between the same two points on the ground. It can be seen from Eq. 1 that the higher the UAV flies, the less ground resolution there will be, using the same image sensor. Similarly, the shorter the focal length, the higher the resolution is. Normally, the resolution and focal length have been selected before a flight. Therefore, Eq. 1 can be used to determine the required flying height to produce the desired

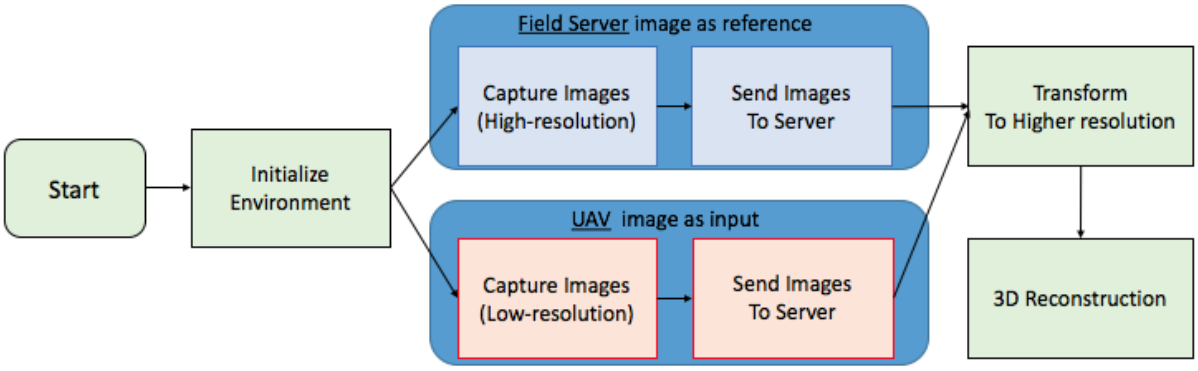

Fig. 2. Proposed System Flowchart

resolution. For example, if the focal length is 4.5 mm, the pixel size is 3.12 µm and the expected resolution is 50 mm, then the flying height will be 72 m.

Once the resolution, flying height and the interested area to be captured have been established, the desired lines of flight and the position of the waypoints can be determined. Flight lines are normally orientated in a north-south or east-west direction and are usually parallel to each other. To collect images for the desired area on the ground, the UAV flies along the entire length of one strip, then move to the next flight line without changing heading and flies backwards along the entire length of the next adjoining flight line. This procedure is repeated until the desired ground area has been completely covered.

The initialization step also need to determine the formation path, and decide whether it use one, two, or more UAVs to take the image. The formation flight is able to optimize the 3D image result because it can take the image in many different angles. The demo of formation flight is illustrated in Fig. 3.

Fig. 3. Flight formation testing using 2 UAVs.

B. Image Acquisition

There are two types of images acquired in the system: from UAV (low-resolution image) and from field server or ground-based device (high-resolution image). For UAV image, there are many components that need to be set such as image overlapping, camera direction, camera angle.

To map a large field with the UAV system, requires a series of images to be taken along each of the multiple flight lines. To guarantee coverage without gaps throughout the field of interest, the images must contain enough overlaps. Aerial image overlap is the amount by which one image includes the area covered by another image, and is expressed as a percentage. Overlap normally contains two types of overlap along two directions: forward overlap and lateral overlap. The forward overlap is the common image area on consecutive images along a flight strip. The lateral overlap encompasses the overlapping areas of images between adjacent flight lines. Fig. 4 illustrates the forward overlap and lateral overlap between two flight lines. In order to cover as much as possible of the area with the minimum number of images taken by the UAV system, it is necessary to investigate appropriate overlap values.

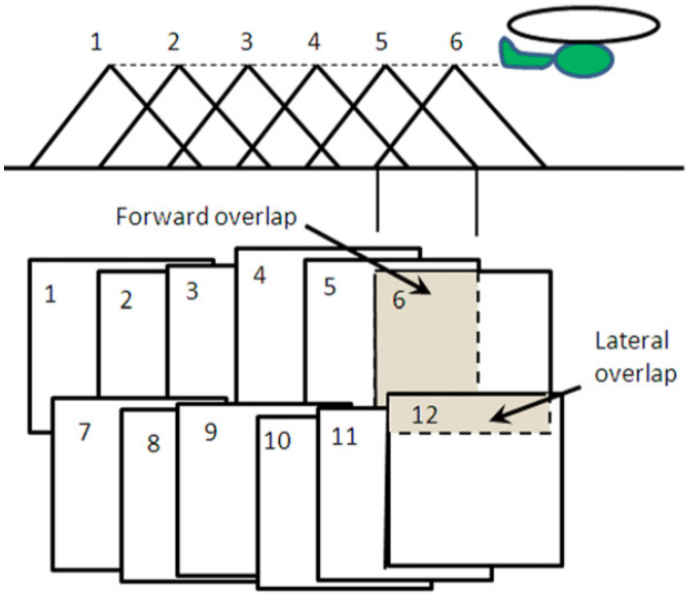

Fig. 4. Aerial image overlapping [14]

To support the formation flight, we also develop automatic shutter for multiple camera which will be installed in each UAV. The illustration is shown in Fig. 5.

Fig. 5. Camera and gimbal which installed in the UAV

C. Super-resolution Technique

The SR algorithm constructs high-resolution images from low-resolution ones. Several methods have been proposed to improve the quality and reduce the computational complexity of SR [15]. The SR algorithm is divided into single and multiple SR. The single SR algorithm constructs an improved image using the parametric image model and prior knowledge/dictionary, whereas the multiple SR algorithm relies on the number of pictures taken as input.

There are many challenges in improving the SR algorithm. In the proposed system, we use a super-resolution algorithm based on adaptive sparse representation via multiple dictionaries for images taken by Unmanned Aerial Vehicles (UAVs) [7]. The super-resolution attainable through the proposed algorithm can increase the precision of 3D reconstruction from UAV images, enabling the production of high-resolution images for constructing high-frequency time series and for high-precision digital mapping in agriculture. The basic idea is to use a field server or ground-based camera to take training images and then construct multiple pairs of dictionaries based on selective sparse representations to reduce instability during the sparse coding process. The dictionaries are classified on the basis of the edge orientation into five clusters: 0, 45, 90, 135, and non-direction. The SR algorithm is expected to reduce blurring, blocking, and ringing artifacts especially in edge areas.

D. 3D Reconstruction

Creating a 3D map with a UAV has become recently popular [16]. A well-known algorithm for this purpose is Structure-from-Motion (SfM) by Furukawa et al. [8]. This algorithm is advantageous to reconstruct and estimate 3D structure from 2D images without any depth information from another sensor such as LiDAR. Using SfM, we only need to rely on the camera and ground control points to construct a 3D image.

The future is wide open for the development of more sophisticated techniques that can recognize known objects or patterns. These advances will help enhance the entire process and make it faster. Currently, we used the GPU, which can modify the process to be even more efficient by replicating image processing tasks and performing many processes in parallel.

III. EXPERIMENTAL RESULTS

To confirm the efficiency of the proposed system, we conducted several experiments. The analysis of these experiments is divided into two subsections: quantitative and qualitative analyses. All experiments were conducted using Matlab R2012b on Win 8.1 64-bit (Intel Core i7@3.2GHz, 8GB). The images used in the experiment were taken at Kazusa DNA Research Institute, Chiba, Japan, red clover tree phenotyping field.

The image dataset consisted of two sub-datasets: training and testing. The training dataset was obtained using a hand-held camera, as shown in Fig. 6. The testing dataset was taken using DJI Phantom 2, as shown in Fig. 7, with an original size of 1280×720 pixels. However, to simplify the process, we divided the image into 320×180 pixel sub-images. In total, we used 300 YCbCr color testing images, as shown in Fig. 8. To conduct the experiment, we enhanced the luminance components (Y) while enhancing the other components using bicubic interpolation. Each resulting image channel was combined to produce a final color image.

Fig. 6. Samples of training images taken by hand-held digital camera.

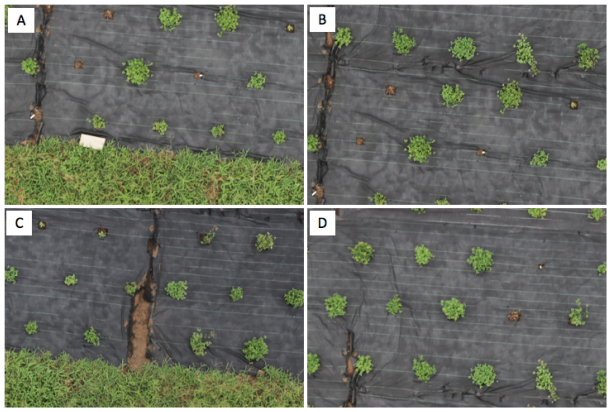

Fig. 7. Images A-D show sample testing images taken by UAV (DJI Phantom 2).

In the experiments, we obtained images by downsampling and blurring the original images and then enlarging using different methods to $3\times$ magnification. We compared the effectiveness of seven methods: nearest neighbor, bilinear, bicubic, Yang et al. [17], Kim et al. [18], Zeyde et al. [19],

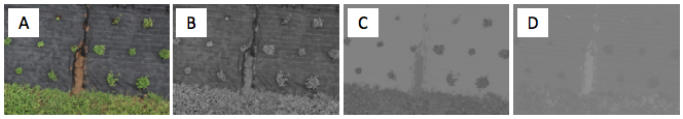

Fig. 8. YCbCr color components. A) Original color image, B) Y component, C) Cb component, D) Cr component.

and the proposed method. The algorithms associated with these vary in nature; therefore, in order to produce objective comparisons, all parameters used in the training and testing had to be similar. However, no specific parameter needed to be used for the conventional interpolation methods.

Our proposed method uses 3×3 patches with no overlapping pixels and five pairs of dictionaries. The algorithm of Yang et al. [17] uses 5×5 patches with a 4-pixels patch overlap and a single pair of dictionaries with 1024 atoms with back-projection. The algorithm of Zeyde et al. [19] uses 3×3 patches with 2-pixels patch overlap and a single pair of dictionaries with 1000 atoms. As mentioned above, these algorithms have different characteristics, and therefore obtaining objective comparisons required that all parameters used in training and testing were similar to those recommended in the respective literature.

A. Quantitative Analysis

Methods for measuring the peak signal-to-noise ration (PSNR) [20], structural similarity (SSIM) [21], feature similarity (FSIM) [22], and elapsed time were used for quantitative measurement. The PSNR in decibels (dB) between the original image and the upscaled image is given by [20]. SSIM is a method that measures the quality of images based on the structural content of the original and magnified images. FSIM is based on the fact that the human visual system processes an image mainly in terms of its low-level features. Two features are considered in FSIM computation: the primary feature, i.e., phase congruency (PC), which is a dimensionless measure of a local structures significance; and the secondary feature, i.e., the image gradient magnitude. FSIM combines both of these features to characterize the local quality of an image. Higher values of PSNR, SSIM, and FSIM indicate better quality. CPU time was computed using Matlab functions (tic and toc) to measure the elapsed time for a certain process. All measurements used only the luminance channel (Y) to simplify and objectively calculate the error.

Table II lists the average values from four measurements, with the best values shown in bold. These result confirm that our proposed method clearly out-performs other methods in terms of PSNR, SSIM, and FSIM. Our method obtains a PSNR value of 25.847 dB, which is at least 11% higher than the other methods. Our proposed method also obtains an SSIM value higher by at least 14% than the other methods. In terms of FSIM, our methods outperforms the others by at least 6%. However, it should be noted that PSNR is not suitable for measuring the quality of bicubic and bilinear, as the quantitative and qualitative analysis for both methods produced some anomalies.

Although our proposed method does not provide the lowest computational time, it is still far better in this respect than Yang et al.'s algorithm [17]. Zeyde et al. [19] produced the lowest computational time in our experiments, while our method competes competitively with Kim et al. [18] with a less than 1 s differential. In future applications and research, the use of a graphics processing unit (GPU) application should offer the opportunity to decrease the computational time of the proposed method.

Nearest neighbor, bilinear, and bicubic were all excluded from the time evaluation as these had salient differences in nature to the proposed and other methods; these conventional methods are simple interpolators that do not use prior information or any learning processes. Moreover, their implementations use Matlab built-in functions, making the comparison unfair as these implement the optimization process automatically.

B. Qualitative Analysis

To evaluate the proposed method in terms of visual results, we conducted experiments using 3× magnified images to compare the proposed method to the other five methods: bilinear, bicubic, Kim et al. [18], Yang et al. [17], and Zeyde et al. [19].

Fig. 9 shows a sample of the experimental results. Our method clearly produces sharper and smoother edges and is able to clearly construct the details of a scene. The other methods all produced images with some artifacts, especially in the line and tree areas, while bicubic and bilinear also produced blurring effects in the enlarged image. Although Yang's and Zeyde's methods generate sharp edge, they still suffer from some noise and produce undesired smoothing. By contrast, Kim's method produces too strong of an edge, with unrealistic results.

Fig. 9 also shows the differences between the original images and the results produces by the respective methods. It is seen that our proposed method has the least amount of difference from the original image, which means that the proposed method has produces the least amount of artifacts, as it can clearly reconstruct edges better than the other algorithms.

IV. APPLICATION TO 3D RECONSTRUCTION

High-resolution imaging is necessary in the construction of high-precision 3D images; correspondingly, the resolution of the input image affects the quality of the 3D reconstruction precision. However, we propose the use of low-resolution image to reduce the payload of the UAV. In this section, we describe an application of our proposed system in 3D reconstruction and then compare it with the other methods.

In this experiment, a DJI Phantom 2 UAV was used to take aerial images of boxes in a field oriented at differing angles, directions, and heights. We use only one UAV because the multiple UAVs are still in the development stage. Before collecting the images, we created a flight plan that considered altitude, latitude, longitude, the distance of each turning point, and flight speed. The UAV periodically collected images

TABLE II
COMPARISON OF THE AVERAGE QUANTITATIVE RESULTS PRODUCED BY PSNR, SSIM, AND FSIM FOR 3× MAGNIFICATION (BOLD FONT INDICATES THE BEST VALUES).

Methods	PSNR	SSIM	FSIM	Time
Nearest neighbor	22.762 ± 3.85	0.637 ± 0.12	0.736 ± 0.06	-
Bilinear	23.243 ± 3.91	0.650 ± 0.12	0.767 ± 0.06	-
Bicubic	23.361 ± 3.93	0.663 ± 0.12	0.779 ± 0.06	-
Kim et al. [18]	23.205 ± 3.93	0.674 ± 0.11	0.789 ± 0.06	5.568 ± 1.83
Yang et al. [17]	23.213 ± 3.93	0.673 ± 0.11	0.795 ± 0.05	67.189 ± 4.78
Zeyde et al. [19]	23.328 ± 3.93	0.677 ± 0.11	0.794 ± 0.05	**0.669** ± 0.04
Proposed	**25.847** ± 4.35	**0.768** ± 0.09	**0.845** ± 0.05	6.290 ± 1.15

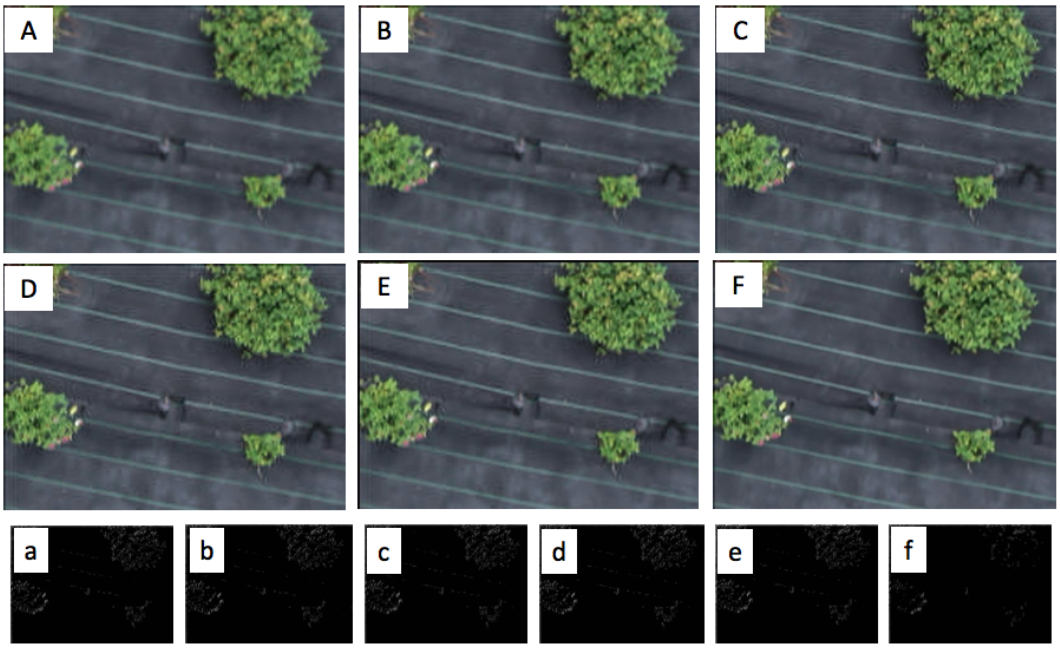

Fig. 9. Results of experiment for 3× magnification (uppercase for color image, lowercase for difference image): A-a) Bilinear, B-b) Bicubic, C-c) Kim et al. [18], D-d) Yang et al. [17], E-e) Zeyde et al. [19], F-f) The proposed method.

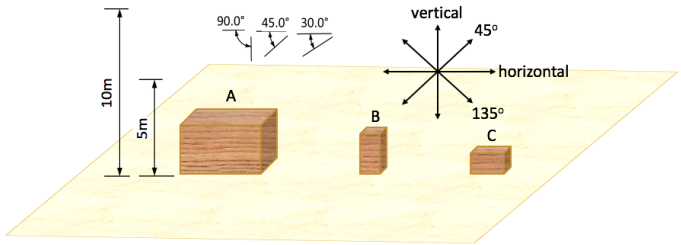

Fig. 10. Flight experimental procedure.

from different angles in order to create 3D images. Only one operator was required to oversee the autonomous flight because the equipment was configured and the UAV can run in fully autonomous mode using defined parameters (e.g., sensors and flight plan). The flight procedure is show in Fig. 10.

The UAV was set to take original image with size 1280 × 720. Then, we created test image by downsampling the original image into size 430 × 260, and later enlarged it using the proposed SR technique. Next, we implemented the Structure from Motion (SfM) algorithm developed by the authors of [8] on a PC (Windows 8.1, 64-bit; CPU: Intel Core i7- 4790, RAM: 32 GB, GPU: GeForce TX780). SfM employs the phenomenon by which humans can recover a 3D structure from a projected 2D (retinal) motion field of a moving object or scene, finding correspondence between images by searching them for features that can be recognized from different angles and distances and matching them in a manner similar to rotating pieces of a jigsaw puzzle until the best match is found. As in a puzzle, the more images that can be sorted and matched, the better the 3D model that can be achieved. The 3D dense cloud image we constructed from various methods is shown in Fig. 11.

Table III lists the result produces by particular methods from matching points with original images. Using a SIFT algorithm, we extracted the feature points from each image and aligned the matching points. The results show that our proposed method produced the highest number of matching points of all of the methods.

Finally, we measured the height and width of each box and then calculated the error by comparing these to the real scale, with the best result indicated by the lowest error value. The height of box A was used as the scale reference to

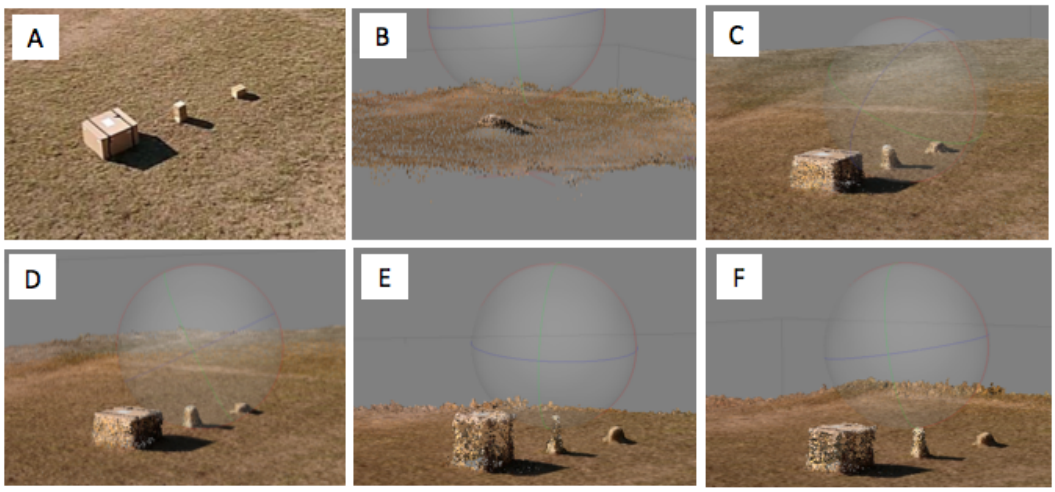

Fig. 11. Results of 3D image reconstruction: (A) Sample of 2D image, (B) From test image, (C) From original image, (D) From the proposed method, (E) From bilinear, (F) From bicubic

TABLE III
RESULTS OF MATCHING POINTS BETWEEN ORIGINAL IMAGES AND PARTICULAR METHODS.

	Camera angle		
Methods	30^o	45^o	90^o
Bilinear	488	976	727
Bicubic	598	1213	844
Proposed method	1812	2708	2068

determine the dimensions of the other boxes. The observed measurements were calculated by taking six sample pairs of points and determining the average distances of each pair using the Euclidian distance measure. The measurement results are listed in Table IV. In the case where the width of box A is 10m, it is seen that our proposed method can decrease the measurement error to a millimeter order of magnitude, while other methods have at least an approximate 11cm error. Some results for the bilinear method could not be calculated owing to bad reconstruction results.

In the case where the imaging was performed from a height of 5m, the original image has the highest precision, even better than the proposed method. However, the proposed method can still keep its measurement error lower than the other methods, and it has the least error in measuring the width of box A.

In the case where imaging occurred from 10m, we found that the proposed system produced an error even lower than that of the original image - a striking result. The greater height of the UAV meant that images with lower detail, or lower amount of pixels per centimeter (PPCM), were produced. In this case, an image taken from 10m has around 1 PPCM, while one taken from 5m has around 2 PPCM. Based on this, we know that the images from 10m suffered at least twice the noise of the 5m image, and the results prove that our proposed method is able to recover test images, reinsert high-frequency details, and repair some of the inconsistency in edges owing to a lowered PPCM.

Bigger, well-shaped objects are easy to reconstruct. In this experiment, we used boxes, not trees, to simplify the experiment. However, in the future we will attempt to conduct real field phenotyping. We note that the lowest error was achieved by our proposed method in calculating the width of box A, which did this with an accuracy within a millimeter order of magnitude. However, for smaller dimensions such as the height of C or the width of B, it will be harder to obtain accurate measurements.

V. CONCLUSIONS

In order to provide more reliable of the agriculture image collection, a UAV based system for high-precision 3D map was studied. The UAV is equipped with a user friendly ground station application which designed to be the interface between a human operator and the UAV to carry out mission planning, flight command activation, and real-time flight monitoring. Based on the navigation data, and the way-points generated by the ground station, the UAV could be automatically navigated to the desired waypoints and hover around each waypoint to collect field image data. By so doing, the aerial images at each point could be captured automatically.

Using only low-resolution, we are able to reduce the payload of the UAV in order to increase the flight time of UAV. The system was tested and proven that it can achieved error of a millimeter order of magnitude. The UAV imaging system developed here offers enhanced capabilities in dealing with agricultural remote sensing needs because, unlike traditional remote sensing systems, it has good spatiotemporal capabilities. When combined with other appropriate sensors like field server, the UAV imaging system has further promise as an agricultural remote sensing platform.

ACKNOWLEDGEMENTS

This work was supported by CREST, Japan Science and Technology Agency. Muhammad Haris would like to thank the Indonesia Endowment Fund for Education (LPDP) Scholarships from the Ministry of Finance, The Republic of Indonesia for doctoral degree scholarship.

TABLE IV

3D MEASUREMENT RESULTS. THE MEASUREMENT IS DETERMINED BY AVERAGING THE DISTANCE OF SIX PAIRS OF POINT'S SAMPLE AND ERROR IS THE DIFFERENCE BETWEEN REAL AND OBSERVED MEASUREMENT (A^* IS USED AS SCALE REFERENCE AND BOLD FONT INDICATE AS THE BEST VALUE.)

			Height (cm)			Width (cm)	
UAV's Height	Methods	A^*	B	C	A	B	C
	Real scale	61	47	25	101	24	32
5m	Original	61	**45.25 (-1.75)**	**22.76 (-2.24)**	106.34 (+5.34)	**22.94 (-1.06)**	**32.27 (+0.27)**
	Proposed	61	42.83 (-4.17)	20.45 (-4.55)	**105.82 (+4.82)**	20.34 (-3.66)	30.26 (-1.74)
	Bicubic	61	41.03 (-5.97)	16.94 (-8.06)	107.54 (+6.54)	20.25 (-3.75)	25.98 (-6.02)
	Bilinear	61	38.93 (-8.07)	17.32 (-7.68)	–	–	–
10m	Original	61	40.71 (-6.29)	20.68 (-4.32)	107.69 (+6.69)	16.73 (-7.27)	26.37 (-5.63)
	Proposed	61	**45.78 (-1.22)**	**28.13 (+3.13)**	**100.37 (-0.63)**	**17.88 (-6.12)**	**28.84 (-3.16)**
	Bicubic	61	40.95 (-6.05)	17.98 (-7.02)	89.73 (-11.27)	12.09 (-11.91)	15.01 (-16.99)
	Bilinear	61	35.18 (-11.82)	–	115.95 (+14.95)	15.06 (-8.94)	–

REFERENCES

[1] S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, "Remote sensing applications for precision agriculture: A learning community approach," *Remote Sensing of Environment*, vol. 88, no. 1, pp. 157–169, 2003.

[2] J. Everaerts *et al.*, "The use of unmanned aerial vehicles (uavs) for remote sensing and mapping," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 1187–1192, 2008.

[3] G. Grenzdörffer, A. Engel, and B. Teichert, "The photogrammetric potential of low-cost uavs in forestry and agriculture," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 31, no. B3, pp. 1207–1214, 2008.

[4] S. C. Chapman, T. Merz, A. Chan, P. Jackway, S. Hrabar, M. F. Dreccer, E. Holland, B. Zheng, T. J. Ling, and J. Jimenez-Berni, "Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping," *Agronomy*, vol. 4, no. 2, pp. 279–301, 2014.

[5] Y. Huang, S. J. Thomson, W. C. Hoffmann, Y. Lan, and B. K. Fritz, "Development and prospect of unmanned aerial vehicle technologies for agricultural production management," *International Journal of Agricultural and Biological Engineering*, vol. 6, no. 3, pp. 1–10, 2013.

[6] C. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: a review," *Precision agriculture*, vol. 13, no. 6, pp. 693–712, 2012.

[7] M. Haris and H. Nobuhara, "Super-resolution based on edge-aware sparse representation via multiple dictionaries," in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 3, 2016, pp. 40–47.

[8] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 8, pp. 1362–1376, 2010.

[9] T. Fukatsu and M. Hirafuji, "Field monitoring using sensor-nodes with a web server," *Journal of Robotics and Mechatronics*, vol. 17, no. 2, pp. 164–172, 2005.

[10] W. Guo, T. Fukatsu, and S. Ninomiya, "Automated characterization of flowering dynamics in rice using field-acquired time-series rgb images," *Plant methods*, vol. 11, no. 1, p. 7, 2015.

[11] T. Sritarapipat, P. Rakwatin, and T. Kasetkasem, "Automatic rice crop height measurement using a field server and digital image processing," *Sensors*, vol. 14, no. 1, pp. 900–926, 2014.

[12] M. Hirafuji, H. Yoichi, T. Kiura, K. Matsumoto, T. Fukatsu, K. Tanaka, Y. Shibuya, A. Itoh, H. Nesumi, N. Hoshi *et al.*, "Creating high-performance/low-cost ambient sensor cloud system using openfs (open field server) for high-throughput phenotyping," in *SICE Annual Conference (SICE), 2011 Proceedings of*. IEEE, 2011, pp. 2090–2092.

[13] D. P. Paine and J. D. Kiser, *Aerial photography and image interpretation*. John Wiley & Sons, 2003.

[14] H. Xiang and L. Tian, "Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (uav)," *Biosystems engineering*, vol. 108, no. 2, pp. 174–190, 2011.

[15] S. C. Park, M. K. Park, and K. M. Gi, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, pp. 21–36, 2003.

[16] F. Remondino, L. Barazzetti, F. Nex, M. Scaioni, and D. Sarazzi, "Uav photogrammetry for mapping and 3d modeling–current status and future perspectives," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, no. 1, p. C22, 2011.

[17] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.

[18] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 6, pp. 1127–1133, 2010.

[19] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*. Springer, 2012, pp. 711–730.

[20] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *Journal of Visual Communication and Image Representation*, vol. 4, no. 4, pp. 324–335, 1993.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[22] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, 2011.

Muhammad Haris received the B.Sc. in Computer Science from the University of Indonesia, Depok, Indonesia, in 2009. He is currently a doctoral student in Department of Intelligent Interaction and Technologies, University of Tsukuba, Japan. His research interests include image processing, super-resolution, image enhancement, and 3D reconstruction especially for aerial images.

Seita Sukisaki received B.Eng from School of Engineering, University of Tsukuba, Japan. He is currently a master student in Department of Intelligent Interaction and Technologies, University of Tsukuba, Japan. His research interests include robotics especially UAV.

Ryo Shimomura received the B.Sc. in Robotics from the Chiba Institute of Technology, Japan in 2016. He is currently a master student in Department of Intelligent Interaction and Technologies, University of Tsukuba, Japan. His research interests include autonomous mobile robot, navigation, point cloud processing, and cooperative control for outdoor-type robot.

Li Hongyang received the B.Sc. in Electronic Information Science and Technology from the Tianjin Normal University, Tianjin, China, in 2010. He is currently research student in Department of Intelligent Interaction and Technologies, University of Tsukuba, Japan. His research interests include UAV aerial image, 3D reconstruction, especially for formation flight.

Zhang Heming recieved B.Eng from School of Electrical Engineering, Nantong University, China. He is now a research student in Department of Intelligent Interaction and Technologies, University of Tsukuba, Japan. His current research include position estimation, position control of UAV, and formation flight.

Hajime Nobuhara received the Dr.Eng. in Computational Intelligence from the Tokyo Institute of Technology, Japan in 2003. He is currently associate Professor at Department of Intelligent Interaction Technologies, University of Tsukuba, Japan. Before that, he was a post doctoral fellow at University of Alberta, Canada. His research interests include computational intelligence, image processing, web intelligence, bio informatics, and UAV.